

RANDOM INTEGRAL MATRICES AND THE COHEN LENSTRA HEURISTICS

MELANIE MATCHETT WOOD

ABSTRACT. We prove that given any $\epsilon > 0$, random integral $n \times n$ matrices with independent entries that lie in any residue class modulo a prime with probability at most $1 - \epsilon$ have cokernels asymptotically (as $n \rightarrow \infty$) distributed as in the distribution on finite abelian groups that Cohen and Lenstra conjecture as the distribution for class groups of imaginary quadratic fields. This is a refinement of a result on the distribution of ranks of random matrices with independent entries in $\mathbb{Z}/p\mathbb{Z}$. This is interesting especially in light of the fact that these class groups are naturally cokernels of square matrices. We also prove the analogue for $n \times (n + u)$ matrices.

1. INTRODUCTION

The Cohen-Lenstra heuristics are conjectures made by Cohen and Lenstra [CL84] on the distribution of class groups of quadratic number fields. For a prime p , we write G_p for the Sylow p -subgroup of an abelian group G . We write $\text{Cl}(K)$ for the class group of a number field K .

Conjecture 1.1 (Cohen and Lenstra [CL84]). *Let p be an odd prime. Let S_X^- be the set of negative fundamental discriminants $D \geq -X$. Let p be an odd prime and B be a finite abelian p -group. Then*

$$\lim_{X \rightarrow \infty} \frac{\#\{D \in S_X^- \mid \text{Cl}(\mathbb{Q}(\sqrt{D})_p \simeq B\}}{|S_X^-|} = \frac{\prod_{k=1}^{\infty} (1 - p^{-k})}{|\text{Aut}(B)|}.$$

Friedman and Washington [FW89] show that if $H_n \in M_{n \times n}(\mathbb{Z}_p)$ is a random matrix drawn with respect to Haar measure on the space of $n \times n$ matrices over \mathbb{Z}_p , then

$$(1) \quad \lim_{n \rightarrow \infty} \mathbb{P}(\text{cok}(H_n) \simeq B) = \frac{\prod_{k=1}^{\infty} (1 - p^{-k})}{|\text{Aut}(B)|}.$$

In other words, cokernels of random p -adic square matrices drawn with respect to Haar measure are distributed according to Cohen and Lenstra's conjectured distribution of class groups (asymptotically as the size of the matrices grows).

Let $K = \mathbb{Q}(\sqrt{D})$ for some $D \in S_X^-$, and S be any finite set of primes of K that generate $\text{Cl}(K)$. We write \mathcal{O}_S^* for the S -units in the integers \mathcal{O}_K , and I_K^S for the abelian group of fractional ideals generated by the elements of S . Then

$$(2) \quad \text{Cl}(K) = \text{cok}(\mathcal{O}_S^* \rightarrow I_K^S),$$

where the map takes $\alpha \mapsto (\alpha)$. So $\text{Cl}(K)_p = \text{cok}(\mathcal{O}_S^* \otimes_{\mathbb{Z}} \mathbb{Z}_p \rightarrow I_K^S \otimes_{\mathbb{Z}} \mathbb{Z}_p)$. Since I_K^S and \mathcal{O}_S^* are both free abelian groups of rank $|S|$ (when $-D > 4$), we have written $\text{Cl}(K)_p$ as a cokernel of a p -adic square matrix $R_D \in M_{n \times n}(\mathbb{Z}_p)$. As D varies, we have a random p -adic square

matrix $R_D \in M_{n \times n}(\mathbb{Z}_p)$ (where n depends on D and we can take $n \rightarrow \infty$) and Conjecture 1.1 is a statement about the distribution of the cokernels of the random matrices R_D .

One might thus imagine that there could be some sense in which the R_D become equidistributed with respect to Haar measure, and that this would imply Conjecture 1.1. However, in this paper we show that in fact having cokernels distributed according to Cohen and Lenstra's conjectured distribution of class groups is a rather robust feature of random matrix regimes, and so much weaker statements (than Haar equidistribution) about the distribution R_D would also imply Conjecture 1.1.

As a particular example, if we take $X_n \in M_{n \times n}(\mathbb{Z}_p)$ whose entries are independent and 0 with probability q and 1 with probability $1 - q$, then (for any q and) for every p , we have Equation (1) with H_n replaced by X_n . These X_n , with their entries concentrated in $\{0, 1\}$, are nowhere near Haar equidistributed in any \mathbb{Z}_p , yet they still have the same cokernel distributions as Haar equidistributed random matrices. More generally, we have the following.

Theorem 1.2. *Let p be a prime and $\epsilon > 0$, and for each n let $X_n \in M_{n \times n}(\mathbb{Z}_p)$ be a random matrix with independent entries. Further, for any entry $(X_n)_{i,j}$ and any $r \in \mathbb{Z}/p\mathbb{Z}$, we require that $\mathbb{P}((X_n)_{i,j} \equiv r \pmod{p}) \leq 1 - \epsilon$. Then for any finite abelian p -group B ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{cok}(X_n) \simeq B) = \frac{\prod_{k=1}^{\infty} (1 - p^{-k})}{|\text{Aut}(B)|}.$$

Note that the matrix entries are not required to be identically distributed and can vary with n . Of course, some condition that the matrix entries are not too concentrated, like $\mathbb{P}((X_n)_{i,j} \equiv r \pmod{p}) \leq 1 - \epsilon$, is certainly necessary, since if the matrices had even two rows whose values were all $r \pmod{p}$, then $\text{cok}(X_n)$ could never be the trivial group.

In fact, in Corollary 3.4, we prove a statement about random integral matrices that implies Theorem 1.2, determining not only the Sylow p -subgroups of their cokernels a single p , but rather the Sylow p -subgroups of their cokernels simultaneously for any finite set of p , and we see (as Cohen and Lenstra [CL84] predict for class groups) that the Sylow p -subgroups for different p are independent.

Of course, the independence of the matrix entries in Theorem 1.2 is a significant hypothesis (and not true in such a form for class groups), and one might wonder to what extent it is necessary. In [Woo14a], it is shown that if one takes the matrices X_n symmetric, but with otherwise independent entries, their cokernels have a *different* distribution than that in Theorem 1.2. The work in that paper was to determine the distribution of Jacobians (a.k.a. sandpile groups) of random graphs, which are a more accessible analogue of class groups. That application also required dealing with the fact that each diagonal entry of the relevant matrix (the graph Laplacian) is dependent on all the entries in its column, and this “small” dependence of the diagonal did not have an effect on the cokernel distribution.

In fact, Cohen and Lenstra [CL84] also make conjectures about class groups of real quadratic (and other totally real abelian) number fields. In particular, if S_X^+ is the set of positive fundamental discriminants $D \leq X$, they conjecture

$$\lim_{X \rightarrow \infty} \frac{\#\{D \in S_X^+ \mid \text{Cl}(\mathbb{Q}(\sqrt{D})_p) \simeq B\}}{|S_X^-|} = \frac{\prod_{k=1}^{\infty} (1 - p^{-k-1})}{|B| |\text{Aut}(B)|}.$$

We see from equation (2) that these class groups are cokernels of $n \times (n+1)$ matrices, since \mathcal{O}_S^* will have rank $|S| + 1$ when the number field K is real quadratic. We in fact prove the following, which follows from Corollary 3.4.

Theorem 1.3. *For any $u \geq 0$, for random $X_n \in M_{n \times (n+u)}(\mathbb{Z}_p)$ with entries as in Theorem 1.2,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{cok}(X_n) \simeq B) = \frac{\prod_{k=1}^{\infty} (1 - p^{-k-u})}{|B|^u |\text{Aut}(B)|}.$$

These distributions on finite abelian groups for other u also arise in the general theory Cohen and Lenstra build to formulate their conjectures.

While the results of this paper are particularly notable for their connection to the Cohen Lenstra heuristics, the proofs in this paper and the history of previous work lie in the fields of additive combinatorics and probability. For $X_n \in M_{n \times n}(\mathbb{Z}_p)$, then $\text{cok}(X_n)$ is trivial if and only if X_n is a non-singular matrix when reduced mod p . More generally, the corank of the matrix mod p is the rank of the cokernel. There is a long history of work on singularity and ranks of the random matrices we consider above mod p , including results of Kozlov [Koz66], Kovalenko and Levitskaja [KL75], Charlap, Rees, and Robbins [CRR90] (first proving Theorem 1.2 in the case that B is the trivial group), Kahn and Komlós [KK01], and Maples [Map10] for general p . However, even our result on ranks (Corollary 3.5), that for X_n as in Theorem 1.2,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{rank}(X_n) = n - k) = p^{-k^2} \prod_{i=1}^k (1 - p^{-i})^{-2} \prod_{i \geq 1} (1 - p^{-i})$$

appears to be new with our hypotheses. The realization that the cokernel distribution, and not just the ranks, should also be insensitive to the distributions of the entries of the matrices is due to Tao and Maples (see [Map13] for some interesting work towards Theorem 1.2).

This robustness of certain statistics of random matrices under changes to the entry distribution of the matrices is an important theme in the study of random matrices and is called *universality* of those statistics. For example, the best upper bounds on the singularity probability of discrete random matrices with independent entries in characteristic 0, due to Bourgain, Vu, and Wood [BVW10], are insensitive to the actual values the entries take (as long as they are not too concentrated).

To prove our main result, we first determine the moments of the cokernel distributions, and from that determine the distributions themselves. Our specific approach was developed in [Woo14a] for the case of symmetric matrices. In this paper, we are able to use a much simplified version of that in [Woo14a] since our matrices have all their entries independent. To find the moments $\mathbb{E}(\# \text{Sur}(\text{cok}(X_n), G))$, we prove inverse Littlewood-Offord theorems (Lemmas 2.1 and 2.7). These both say that if values of several linear functions of our n independent variables are not close to equidistributed, then the linear functions are close to having extra structure. The extra structure is analogous to having a linear dependence in rows of a matrix after deleting a small number of columns, but since our linear algebra is not always over a field there are many layers to the type of dependence we can have, which are captured by our notion of *depth*. Inverse Littlewood-Offord Theorems are a key component in the most recent work on singularity probability of discrete random matrices in characteristic 0, both [BVW10] mentioned above and the earlier work of Tao and Vu [TV07].

(See the papers of Tao and Vu [TV10] and Nguyen and Vu for the most recent [NV11] inverse Littlewood-Offord Theorems in characteristic 0, as well as a guide to the extensive previous work on the problem.) However, there are significant differences in the actual mathematics of these theorems in characteristic 0 versus characteristic p , since in characteristic 0 one doesn't expect any kind of equidistribution, but rather just a good upper bound on the probabilities. Maples [Map13] proves a Littlewood-Offord Theorem in characteristic p that is not strong enough for our purposes for fixed p , but does have the advantage of uniformity in p .

To finally determine our cokernel distribution from the moments, we can't rely on the usual probabilistic methods such as Carleman's condition (since our moments are too big—our k th moment is of order $p^{k^2/2}$). However, we use a specifically tailored result [Woo14a] that in our cases shows that that moments we obtain determine a unique distribution. This situation of needing to show fast growing moments of random abelian groups determine the distribution of the groups has arisen before in number theory, both in Cohen-Lenstra problems, e.g. in the work of Fouvry and Klüners [FK06] and Ellenberg, Venkatesh, and Westerland [EVW09], and in a related problem about Selmer groups in work of Heath-Brown [HB94]. Ellenberg, Venkatesh, and Westerland [EVW09] make progress towards proving the function field analogue of the Cohen-Lenstra heuristics by proving new homological stability theorems that determine some of the moments of the relevant class groups.

1.1. Further notation. We use $[n]$ to denote $\{1, \dots, n\}$. We write $\text{Hom}(A, B)$ and $\text{Sur}(A, B)$ for the set of homomorphisms and surjections, respectively, from A to B . We write \mathbb{P} for probability and \mathbb{E} for expected value. We write $\exp(x)$ for the exponential function e^x .

Throughout the paper, we let a be a positive integer and let $R = \mathbb{Z}/a\mathbb{Z}$. We then study finite abelian groups G whose exponent divides a , i.e. $aG = 0$. We write G^* for $\text{Hom}(G, R)$.

2. FINDING THE MOMENTS

We will study integral matrices by reducing them mod a for all a (and analogously matrices over \mathbb{Z}_p by reducing them mod p^k for all k). For the rest of the paper, let a be a positive integer and let $R = \mathbb{Z}/a\mathbb{Z}$. We change the notation for our random matrix from the introduction to make it easier to read our proof. Let M be a random $n \times (n + u)$ matrix with entries in R . We let $M_1, \dots, M_{n+u} \in R^n$ be the columns of M , and m_{ij} the entries of M (so that the entries of M_j are m_{ij}). We let $\epsilon > 0$.

The following definition captures the two hypotheses of Theorem 1.2: independence of entries and entries not too concentrated.

Definition. A random variable y in a ring T is ϵ -balanced if for every maximal ideal \wp of T and $r \in T/\wp$ we have $\mathbb{P}(y \equiv r \pmod{\wp}) \leq 1 - \epsilon$ (e.g., $y \in R$ is ϵ -balanced if for every prime $p \mid a$ and $r \in \mathbb{Z}/p\mathbb{Z}$ we have $\mathbb{P}(y \equiv r \pmod{p}) \leq 1 - \epsilon$). A random vector or matrix is ϵ -balanced if its entries are independent and ϵ -balanced

We let $V = R^n$ with basis v_i and $W = R^m$ with basis w_j . Note for $\sigma \subset [n]$, V has distinguished submodules $V_{\setminus \sigma}$ generated by the v_i with $i \notin \sigma$. (So $V_{\setminus \sigma}$ comes from not using the σ coordinates.) We view $M \in \text{Hom}(W, V)$ and its columns M_j as elements of V so that $M_j = Mw_j = \sum_i m_{ij}v_i$. Let G be a finite abelian group with exponent dividing a . We have $\text{cok } M = V/MW$. To investigate the moments $\mathbb{E}(\#\text{Sur}(\text{cok } M, G))$, we recognize that each

such surjection lifts to a surjection $V \rightarrow G$ and so we have

$$(3) \quad \mathbb{E}(\#\text{Sur}(\text{cok } M, G)) = \sum_{F \in \text{Sur}(V, G)} \mathbb{P}(F(MW) = 0).$$

If M is ϵ -balanced, then by the independence of columns, we have

$$\mathbb{P}(F(MW) = 0) = \prod_{j=1}^m \mathbb{P}(F(M_j) = 0).$$

So we aim to estimate these probabilities $\mathbb{P}(F(M_j) = 0)$. We will first estimate these for the vast majority of F , which satisfy the following helpful property.

Definition. We say that $F \in \text{Hom}(V, G)$ is a *code* of distance w , if for every $\sigma \subset [n]$ with $|\sigma| < w$, we have $FV_{\setminus \sigma} = G$. In other words, F is not only surjective, but would still be surjective if we throw out (any) fewer than w of the standard basis vectors from V . (If a is prime so that R is a field, then this is equivalent to whether the transpose map $F : G^* \rightarrow V^*$ is injective and has image $\text{im}(F) \subset V^*$ a linear code of distance w , in the usual sense.)

Lemma 2.1. *Let R and G be as above. Let $\epsilon > 0$ and $\delta > 0$. Let X be an ϵ -balanced random vector in V . Let $F \in \text{Hom}(V, G)$ be a code of distance δn and $A \in G$. For all n we have*

$$|\mathbb{P}(FX = A) - |G|^{-1}| \leq \exp(-\epsilon \delta n / a^2).$$

Let ζ be a primitive a th root of unity. To prove Lemma 2.1, we will use the discrete Fourier transform and the following basic estimate.

Lemma 2.2. *Let y be an entry of an ϵ -balanced random variable in R , and let m be an integer such that $\zeta^m \neq 1$. Then $|\mathbb{E}(\zeta^{my})| \leq \exp(-\epsilon/a^2)$.*

Proof. This is proven in [Woo14b, Proof of Lemma 4.1]. Briefly, the longest $|\mathbb{E}(\zeta^y)|$ could be was if ζ^y was one a th root of unity $1 - \epsilon$ of the time, and a consecutive (around the unit circle) a th root of unity the rest of the time. \square

Proof of Lemma 2.1. We have, by the discrete Fourier transform,

$$\mathbb{P}(FX = A) = |G|^{-1} \sum_{C \in G^*} \mathbb{E}(\zeta^{C(FX-A)}) = |G|^{-1} + |G|^{-1} \sum_{C \in G^* \setminus \{0\}} \mathbb{E}(\zeta^{C(-A)}) \prod_{1 \leq i \leq n} \mathbb{E}(\zeta^{C(v_i)X_i}).$$

Since $C \neq 0$ and F is a code, there must be at least δn values of i such that $F(v_i) \notin \ker C$. So using Lemma 2.2, we have

$$|\mathbb{P}(FX = A) - |G|^{-1}| = \left| \mathbb{E}(\zeta^{C(-A)}) \prod_{1 \leq k \leq n} \mathbb{E}(\zeta^{C(v_i)X_i}) \right| \leq \exp(-\epsilon \delta n / a^2).$$

\square

We then put these estimates for columns together using a simple inequality.

Lemma 2.3. *If we have integer $m \geq 2$ and real numbers $x \geq 0$ and y such that $|y|/x \leq 2^{1/(m-1)} - 1$ and $x + y \geq 0$, then*

$$x^m - 2mx^{m-1}|y| \leq (x + y)^m \leq x^m + 2mx^{m-1}|y|.$$

Proof. We can assume $x = 1$ by homogeneity. We divide into two cases based on the sign of y . Then note the middle and right hand expressions are equal when $y = 0$ and the derivative of the middle is at most the derivative of the right when $0 \leq y \leq 2^{1/(m-1)} - 1$. A similar argument when $-1 \leq y \leq 0$ compares the left and middle expressions. \square

Lemma 2.4. *Let R , G , and u be as above. Let $\epsilon > 0$ and $\delta > 0$. Then there are $c, K > 0$ such that the following holds. Let $M \in \text{Hom}(W, V)$ be ϵ -balanced random matrix. Let $F \in \text{Hom}(V, G)$ be a code of distance δn . Let $A \in \text{Hom}(W, G)$. For all n we have*

$$|\mathbb{P}(FM = A) - |G|^{-n-u}| \leq \frac{K \exp(-cn)}{|G|^{n+u}}.$$

Proof. For n large enough, we have

$$\exp(-\epsilon \delta n / a^2) |G| \leq \log 2 / (n + u - 1) \leq 2^{1/(n+u-1)} - 1,$$

since $2^{1/(n+u-1)} - 1 = e^{\log 2 / (n+u-1)} - 1 \geq \log 2 / (n + u - 1)$. So for n sufficiently large, we can combine Lemma 2.1 and Lemma 2.2 to obtain

$$|\mathbb{P}(FM = A) - |G|^{-n-u}| \leq 2(n + u) \exp(-\epsilon \delta n / a^2) |G|^{-n-u+1}.$$

The lemma follows. \square

So far, we have dealt with $F \in \text{Hom}(V, G)$ that are codes. Unfortunately, it is not sufficient to divide F into codes and non-codes. We need a more delicate division of F based on the subgroups of G . This division can be approximately understood as separating the F based on what largest size subgroup they are a code for. For an integer D with prime factorization $\prod_i p_i^{e_i}$, let $\ell(D) = \sum_i e_i$. The following concept was introduced in [Woo14b]. Since $V_{\setminus \sigma}$ is a subgroup of V , for $F \in \text{Hom}(V, G)$, the image $F(V_{\setminus \sigma})$ is a subgroup of G .

Definition. The *depth* of an $F \in \text{Hom}(V, G)$ is the maximal positive D such that there is a $\sigma \subset [n]$ with $|\sigma| < \ell(D) \delta n$ such that $D = [G : F(V_{\setminus \sigma})]$, or is 1 if there is no such D .

Remark 2.5. In particular, if the depth of F is 1, then for every $\sigma \subset [n]$ with $|\sigma| < \delta n$, we have that $F(V_{\setminus \sigma}) = G$ (as otherwise $\ell([G : F(V_{\setminus \sigma})]) \geq 1$), and so we see that F is a code of distance δn .

We have a bound the number of F that we have of depth D .

Lemma 2.6 (Count F of given depth, Lemma 5.2 of [Woo14b]). *There is a constant K depending on G such that if $D > 1$, then number of $F \in \text{Hom}(V, G)$ of depth D is at most*

$$K \binom{n}{\lceil \ell(D) \delta n \rceil - 1} |G|^n |D|^{-n + \ell(D) \delta n}.$$

Now for each depth, we will get a bound on $\mathbb{P}(FM = 0)$, with the smaller the depth, the better the bound.

Lemma 2.7 (Bound probability for column given depth). *Let R, G be as above. Let $\epsilon > 0$ and $\delta > 0$. If $F \in \text{Hom}(V, G)$ has depth $D > 1$ and $[G : F(V)] < D$, then for all ϵ -balanced random vectors X in V and all n ,*

$$\mathbb{P}(FX = 0) \leq (1 - \epsilon) (D |G|^{-1} + \exp(-\epsilon \delta n / a^2)).$$

Proof. Pick a $\sigma \subset [n]$ with $|\sigma| < \ell(D)\delta n$ such that $D = [G : F(V_{\setminus\sigma})]$. Let $F(V_{\setminus\sigma}) = H$. However, since $[G : F(V)] < D$, we cannot have σ empty. We have $FX = \sum_{i \notin \sigma} F(v_i)X_i + \sum_{i \in \sigma} F(v_i)X_i$. So

$$\mathbb{P}(FX = 0) = \mathbb{P}\left(\sum_{i \in \sigma} F(v_i)X_i \in H\right) \mathbb{P}\left(\sum_{i \notin \sigma} F(v_i)X_i = -\sum_{i \in \sigma} F(v_i)X_i \mid \sum_{i \in \sigma} F(v_i)X_i \in H\right).$$

For the first factor, we note that since $[G : F(V)] < D$, there must be some $i \in \sigma$ with the reduction $F(v_i) \neq 0 \in G/H$. Thus conditioning on all other X_k for $k \neq i$, by the ϵ -balanced assumption on X , we have that $\mathbb{P}(\sum_{i \in \tau} F(v_i)X_i \in H) \leq 1 - \epsilon$.

Then, we note that the restriction of F to $V_{\setminus\sigma}$ is a code of distance δn in $\text{Hom}(V_{\setminus\sigma}, H)$. (If it were not, then by eliminating σ and $< \delta n$ indices, we would eliminate $< (\ell(D) + 1)\delta n$ indices and have an image which was index that D strictly divides, contradicting the depth of F .) So conditioning on the X_i with $i \in \sigma$, we can estimate the conditional probability above using Lemma 2.1:

$$\mathbb{P}\left(\sum_{i \notin \sigma} F(v_i)X_i = -\sum_{i \in \sigma} F(v_i)X_i \mid \sum_{i \in \sigma} F(v_i)X_i \in H\right) \leq |H|^{-1} + \exp(-\epsilon\delta n/a^2).$$

The lemma follows. \square

Lemma 2.8 (Bound probability for matrix given depth). *Let R, G, u be as above. Let $\epsilon > 0$ and $\delta > 0$. Then there is a real K such that if $F \in \text{Hom}(V, G)$ has depth $D > 1$ and $[G : F(V)] < D$ (e.g. the latter is true if $F(V) = G$), then for all ϵ -balanced random matrices $M \in \text{Hom}(W, V)$, and all n ,*

$$\mathbb{P}(FM = 0) \leq K \exp(-\epsilon n) D^n |G|^{-n}.$$

Proof. By the independence of the columns of M , we can take the $n + u$ th power of the bound in Lemma 2.7, and apply Lemma 2.3. We have, for n large enough

$$\begin{aligned} & (1 - \epsilon)^{n+u} (D|G|^{-1} + \exp(-\epsilon\delta n/a^2))^{n+u} \\ & \leq \exp(-\epsilon(n+u)) (D^{n+u}|G|^{-n-u} + 2(n+u) \exp(-\epsilon\delta n/a^2) D^{n+u-1}|G|^{-n-u+1}). \end{aligned}$$

The lemma follows. \square

Now we can combine the estimates we have for $\mathbb{P}(FM = 0)$ for F of various depth with the bounds we have on the number of F of each depth to obtain our main result on the moments of cokernels of random matrices.

Theorem 2.9. *Let a be a positive integer, and u be a non-negative integer. Let $\epsilon > 0$ be a real number and G a finite abelian group with exponent dividing a . Then there are $c, K > 0$ such that the following holds. Let M be an ϵ -balanced $n \times (n+u)$ random matrix with entries in $\mathbb{Z}/a\mathbb{Z}$.*

$$|\mathbb{E}(\#\text{Sur}(\text{cok}(M), G)) - |G|^{-u}| \leq K e^{-cn}.$$

Proof. By Equation (3), we need to estimate $\sum_{F \in \text{Sur}(V, G)} \mathbb{P}(FM = 0)$. We let K change in each line, as long as it is a constant depending only on a, u, ϵ, G . Take $d < \min(\epsilon, \log(2))$.

Using Lemmas 2.6 and 2.8 we have

$$\begin{aligned}
\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} \mathbb{P}(FX = 0) &\leq \sum_{\substack{D > 1 \\ D \mid \#G}} \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ depth } D}} \mathbb{P}(FX = 0) \\
&\leq \sum_{\substack{D > 1 \\ D \mid \#G}} K \binom{n}{\lceil \ell(D) \delta n \rceil - 1} |G|^n D^{-n + \ell(D) \delta n} \exp(-\epsilon n) D^n |G|^{-n} \\
&\leq \sum_{\substack{D > 1 \\ D \mid \#G}} K \binom{n}{\lceil \ell(D) \delta n \rceil - 1} D^{\ell(D) \delta n} \exp(-\epsilon n) \\
&\leq K \binom{n}{\lceil \ell(|G|) \delta n \rceil - 1} |G|^{\ell(|G|) \delta n} \exp(-\epsilon n) \\
&\leq K e^{-dn},
\end{aligned}$$

as long as we choose δ small enough.

Also, from Lemma 2.6, we can choose δ small enough so that we have

$$\begin{aligned}
\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} |G|^{-n-u} &\leq \sum_{\substack{D > 1 \\ D \mid \#G}} \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ depth } D}} |G|^{-n-u} \\
&\leq \sum_{\substack{D > 1 \\ D \mid \#G}} K \binom{n}{\lceil \ell(D) \delta n \rceil - 1} |G|^n |D|^{-n + \ell(D) \delta n} |G|^{-n} \\
&\leq K \binom{n}{\lceil \ell(|G|) \delta n \rceil - 1} 2^{-n + \ell(|G|) \delta n} \\
&\leq K e^{-dn}.
\end{aligned}$$

We also have

$$\begin{aligned}
\sum_{F \in \text{Hom}(V, G) \setminus \text{Sur}(V, G)} |G|^{-n-u} &\leq \sum_{H \text{ proper s.g. of } G} \sum_{F \in \text{Hom}(V, H)} |G|^{-n-u} \\
&\leq \sum_{H \text{ proper s.g. of } G} |H|^{n+u} |G|^{-n-u} \\
&\leq K e^{-dn}.
\end{aligned}$$

Then given a choice of δ that satisfies the two requirements above, using Lemma 2.4 we have a c such that

$$\sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ code of distance } \delta n}} |\mathbb{P}(FX = 0) - |G|^{-m}| \leq K e^{-cn}.$$

If necessary, we take c smaller so $c \leq d$. In conclusion,

$$\begin{aligned}
& \left| \left(\sum_{F \in \text{Sur}(V, G)} \mathbb{P}(FX = 0) \right) - |G|^{-u} \right| = \left| \left(\sum_{F \in \text{Sur}(V, G)} \mathbb{P}(FX = 0) \right) - \left(\sum_{F \in \text{Hom}(V, G)} |G|^{-n-u} \right) \right| \\
& \leq \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ code of distance } \delta n}} |\mathbb{P}(FX = 0) - |G|^{-n-u}| + \sum_{\substack{F \in \text{Sur}(V, G) \\ F \text{ not code of distance } \delta n}} \mathbb{P}(FX = 0) + \sum_{\substack{F \in \text{Hom}(V, G) \\ F \text{ not code of dist. } \delta n}} |G|^{-n-u} \\
& \leq Ke^{-cn}.
\end{aligned}$$

□

3. MOMENTS DETERMINE THE DISTRIBUTION

We use the following theorem to determine the asymptotic distribution of $\text{cok}(M)$ as $n \rightarrow \infty$ from the moments in Theorem 2.9.

Theorem 3.1 (c.f. Theorem 8.3 in [Woo14b]). *Let X_n and Y_n be sequences of random finitely generated abelian groups. Let a be a positive integer and A be the set of (isomorphism classes of) abelian groups with exponent dividing a . Suppose that for every $G \in A$, we have a number $M_G \leq |\wedge^2 G|$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{E}(\# \text{Sur}(X_n, G)) = M_G.$$

Then for every $H \in A$, the limit $\lim_{n \rightarrow \infty} \mathbb{P}(X_n \otimes \mathbb{Z}/a\mathbb{Z} \simeq H)$ exists, and for all $G \in A$ we have

$$\sum_{H \in A} \lim_{n \rightarrow \infty} \mathbb{P}(X_n \otimes \mathbb{Z}/a\mathbb{Z} \simeq H) \# \text{Sur}(H, G) = M_G.$$

If for every $G \in A$, we also have $\lim_{n \rightarrow \infty} \mathbb{E}(\# \text{Sur}(Y_n, G)) = M_G$, then, we have that for every every $H \in A$

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \otimes \mathbb{Z}/a\mathbb{Z} \simeq H) = \lim_{n \rightarrow \infty} \mathbb{P}(Y_n \otimes \mathbb{Z}/a\mathbb{Z} \simeq H).$$

For the rest of this section, we fix a non-negative integer u . We construct a random abelian group according to Cohen and Lenstra's distribution for each u as follows. Let P be the set of primes dividing a . Independently for each p , we have a random finite abelian p -group Y_p given by taking each group B with probability

$$\frac{\prod_{k=1}^{\infty} (1 - p^{-k-u})}{|B|^u |\text{Aut}(B)|}.$$

We then form a random group Y by taking $Y = \prod_{p \in P} Y_p$.

Lemma 3.2. *For every finite abelian group G with exponent dividing a , we have*

$$\mathbb{E}(\# \text{Sur}(Y, G)) = |G|^{-u}.$$

In particular, taking G the trivial group says that the above distribution on B is a probability distribution.

Proof. By factoring over primes $p \in P$, we can reduce to the case when $P = \{p\}$. Let \mathcal{A} be the set of finite abelian p -groups. Multiplying [CL84, Proposition 4.1 (ii)] (for $k = \infty$ and $K = A$) by $|\text{Aut}(K)|$, we obtain, for every i ,

$$\sum_{B \in \mathcal{A}, |B|=p^i} \frac{|\text{Sur}(B, G)|}{|\text{Aut}(B)|} = \sum_{B \in \mathcal{A}, |B|=p^i/|G|} \frac{1}{|\text{Aut}(B)|}.$$

Dividing by p^{iu} and summing over all i , we obtain

$$\sum_{B \in \mathcal{A}} \frac{|\text{Sur}(B, G)|}{|B|^u |\text{Aut}(B)|} = |G|^{-u} \sum_{B \in \mathcal{A}} \frac{1}{|B|^u |\text{Aut}(B)|}.$$

By [CL84, Corollary 3.7 (i)] (with $s = u$ and $k = \infty$), we have $\sum_{B \in \mathcal{A}} |B|^{-u} |\text{Aut}(B)|^{-1} = \prod_{j \geq 1} (1 - p^{-j-u})^{-1}$, and the lemma follows. \square

We can now determine the distribution of our cokernels by comparing their moments to those of Y .

Corollary 3.3 (of Theorem 2.9 and Theorem 3.1). *Let $\epsilon > 0$ and let $M \in M_{n \times (n+u)}(\mathbb{Z})$ (resp, $M \in M_{n \times (n+u)}(\mathbb{Z}_p)$) be an ϵ -balanced random matrix. Let G be a finite abelian group with exponent dividing a (resp., p^k). For Y defined above,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{cok}(M) \otimes \mathbb{Z}/a\mathbb{Z} \simeq G) = \mathbb{P}(\text{cok}(Y) \otimes \mathbb{Z}/a\mathbb{Z} \simeq G).$$

In particular, we can conclude the following, which proves Theorems 1.2 and 1.3.

Corollary 3.4. *Let $\epsilon > 0$ and let $M \in M_{n \times (n+u)}(\mathbb{Z})$ (resp, $M \in M_{n \times (n+u)}(\mathbb{Z}_p)$) be an ϵ -balanced random matrix. Let B be a finite abelian group (resp., finite abelian p -group). Let P be a finite set of primes including all those dividing $|B|$ (resp., $P = \{p\}$). Let $H_P := \prod_{p \in P} H_p$. Then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{cok}(M)_P \simeq B) = \frac{1}{|B|^u |\text{Aut}(B)|} \prod_{p \in P} \prod_{k=1}^{\infty} (1 - p^{-k-u}).$$

Proof. Note that if B is a finite abelian group with exponent that has prime factorization $\prod_{p \in P} p^{e_p}$, then if we take $a = \prod_{p \in P} p^{e_p+1}$, for any finitely generated abelian group H , we have $H \otimes \mathbb{Z}/a\mathbb{Z} \simeq G$ if and only if $H_P \simeq G$.

So the corollary follows from Corollary 3.3 and the construction of Y \square

Also taking $a = p$ for a prime p in Corollary 3.3, we conclude the following on the distribution of p -ranks.

Corollary 3.5. *Let p be a prime and $\epsilon > 0$. Let $\epsilon > 0$ and let $M \in M_{n \times (n+u)}(\mathbb{Z}/p\mathbb{Z})$ be an ϵ -balanced random matrix. For every non-negative integer k*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\text{rank}(M) = n - k) = p^{-k(k+u)} \prod_{i=1}^k (1 - p^{-i})^{-1} \prod_{i=1}^{k+u} (1 - p^{-i})^{-1} \prod_{i \geq 1} (1 - p^{-i})$$

Proof. We apply Theorem 2.9 with $a = p$ and Theorem 3.1 with $a = p$ to $\text{cok}(M)$ and Y . We can read off the rank distribution of Y from [CL84][Theorem 6.3]. (Alternatively, instead of Y we could use cokernels of $H_n \in M_{n \times (n+u)}(\mathbb{Z}/p\mathbb{Z})$ from the uniform distribution and use the elementary count of matrices over $\mathbb{Z}/p\mathbb{Z}$ of a given rank.) \square

Acknowledgements. This work was done with the support of an American Institute of Mathematics Five-Year Fellowship and National Science Foundation grants DMS-1301690. The author thanks John Voight for useful comments on an earlier draft of this paper.

REFERENCES

- [BVW10] Jean Bourgain, Van H. Vu, and Philip Matchett Wood. On the singularity probability of discrete random matrices. *Journal of Functional Analysis*, 258(2):559–603, 2010.
- [CL84] H. Cohen and H. W. Lenstra, Jr. Heuristics on class groups of number fields. In *Number theory, Noordwijkerhout 1983 (Noordwijkerhout, 1983)*, volume 1068 of *Lecture Notes in Math.*, pages 33–62. Springer, Berlin, 1984.
- [CRR90] Leonard S. Charlap, Howard D. Rees, and David P. Robbins. The asymptotic probability that a random biased matrix is invertible. *Discrete Mathematics*, 82(2):153–163, June 1990.
- [EVW09] Jordan S. Ellenberg, Akshay Venkatesh, and Craig Westerland. Homological stability for Hurwitz spaces and the Cohen-Lenstra conjecture over function fields. *arXiv:0912.0325 [math]*, December 2009.
- [FK06] Étienne Fouvry and Jürgen Klüners. Cohen–Lenstra Heuristics of Quadratic Number Fields. In Florian Hess, Sebastian Pauli, and Michael Pohst, editors, *Algorithmic Number Theory*, number 4076 in *Lecture Notes in Computer Science*, pages 40–55. Springer Berlin Heidelberg, January 2006.
- [FW89] Eduardo Friedman and Lawrence C. Washington. On the distribution of divisor class groups of curves over a finite field. In *Théorie des nombres (Quebec, PQ, 1987)*, pages 227–239. de Gruyter, Berlin, 1989.
- [HB94] D. R. Heath-Brown. The size of Selmer groups for the congruent number problem. II. *Inventiones Mathematicae*, 118(2):331–370, 1994. With an appendix by P. Monsky.
- [KK01] Jeff Kahn and JÁnos KomlÓs. Singularity Probabilities for Random Matrices over Finite Fields. *Combinatorics, Probability and Computing*, 10(02):137–157, March 2001.
- [KL75] I. N. Kovalenko and A. A. Levitskaja. Limiting behavior of the number of solutions of a system of random linear equations over a finite field and a finite ring. *Doklady Akademii Nauk SSSR*, 221(4):778–781, 1975.
- [Koz66] M. V. Kozlov. On the rank of matrices with random Boolean elements. *Soviet Mathematics. Doklady*, 7:1048–1051, 1966.
- [Map10] Kenneth Maples. Singularity of Random Matrices over Finite Fields. *arXiv:1012.2372 [math]*, December 2010.
- [Map13] Kenneth Maples. Cokernels of random matrices satisfy the Cohen-Lenstra heuristics. *arXiv:1301.1239 [math]*, January 2013.
- [NV11] Hoi Nguyen and Van Vu. Optimal inverse Littlewood–Offord theorems. *Advances in Mathematics*, 226(6):5298–5319, April 2011.
- [TV07] Terence Tao and Van Vu. On the singularity probability of random Bernoulli matrices. *Journal of the American Mathematical Society*, 20(3):603–628, 2007.
- [TV10] Terence Tao and Van Vu. A sharp inverse Littlewood–Offord theorem. *Random Structures & Algorithms*, 37(4):525–539, 2010.
- [Woo14a] Melanie Matchett Wood. The distribution of sandpile groups of random graphs. *arXiv:1402.5149 [math]*, February 2014.
- [Woo14b] Melanie Matchett Wood. Parametrization of ideal classes in rings associated to binary forms. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2014(689):169–199, April 2014.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, 480 LINCOLN DRIVE, MADISON, WI 53705 USA, AND AMERICAN INSTITUTE OF MATHEMATICS, 600 EAST BROKAW ROAD, SAN JOSE, CA 95112 USA

E-mail address: mmwood@math.wisc.edu